

PigQTLdb - A Pig QTL Database

Zhi-Liang Hu¹, Svetlana Dracheva², Wonhee Jang², Donna Maglott², John Bastiaansen³, James Reecy¹, Max Rothschild¹

¹ Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University; ² National Center for Biotechnology Information; ³ Sygen International

Abstract

Mapping efforts of quantitative trait loci (QTL) in pigs during the past decade have resulted in hundreds of QTL reported for growth, meat quality, reproduction, disease resistance and other traits. It is a challenge to correctly locate, interpret and compare QTL results from different studies. We developed a relational database (PigQTLdb) integrate all available pig QTL data in the public domain and thus facilitate the use of QTL data in further studies. We also developed a "trait ontology" to standardize names of traits and to simplify organization of the data. These steps made it possible to compare primary data from diverse sources and methods. Efforts were also made to utilize existing pig map databases and other publicly available data resources (such as PubMed) to avoid redundant developmental work. The PigQTLdb was also designed to include data representing major genes and markers associated with large effect on economically important traits. To date, over 790 QTLs from 78 publications have been curated into the database. Those QTLs cover more than 300 different traits. These data have been submitted to the Gene and Map Viewer resources at NCBI, where the information about markers has been matched to marker records in NCBI's UniSTS database. This allows automatic matching of markers to public sequence data by e-PCR. The data are retrievable from NCBI via Gene, Map Viewer, and UniSTS. All efforts were undertaken to improve integrated functional genomics resources for pigs.

Introduction

The ultimate goal of QTL studies is the identification of the gene(s) responsible for the phenotypic variation observed in a particular trait. However, it is a challenging task to combine results from different research projects involving QTL studies for positional candidate gene or marker searches from information scattered in many publications, each described in different ways (Bidel and Rothschild, 2002). The pig QTL database we developed can be used for easy search and comparison of QTL results from different studies, derived from different populations, and obtained with perhaps different testing methods. With overlapping QTL results from different studies, it will be easier to confirm and narrow down QTL regions that may help to speed up the positional search for underlying genes.

Material and Methods

All QTL data are from published journal papers and reports. Over 50 parameters / data types are subjected for raw data collection. These include: QTL location (Chromosome, Location Span), Flanking markers (at suggestive and / or significant test levels), Peak markers, Test Statistics (LOD score, Least square means, P-values, F-values, Variance), QTL effects (Dominance effect, Additive effect), Candidate genes, Traits and their measurements, etc. The data are organized into related tables (Figure 1) and loaded into MySQL, a relational database software.

The programs to make worldwide web user interface to the data is written in Perl / CGI. Communication between the database and the web server was made possible with Perl DBI / DBD. Lincoln Stein's Perl GD library was employed to draw QTL map graphs "on the fly". All data and tools are hosted on a RedHat Linux server located at Iowa State University.

The QTL pipeline at the National Center for Biotechnology Information (NCBI) was utilized to facilitate the housing of pig QTL data at NCBI. The mapping data for chromosomal maps was mainly from USDA-MARC (Rohrer et al., 1996).

Figure 1. A conceptual database schema showing how the pig QTL data are organized and linked out.

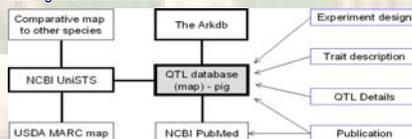
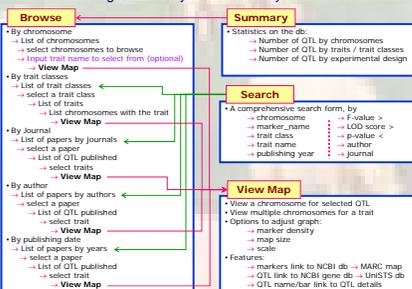


Figure 2. A schematic diagram showing different ways how information in PigQTLdb may be accessed by search and browse.



Results

To date, 791 QTLs from 73 publications have been curated into the database. Those QTLs represent 219 different traits. The PigQTLdb is implemented at both NCBI and on NAGRP Animal Genome server at Iowa State University (URL: <http://www.animalgenome.org/QTLdb/>), each with some unique features to facilitate user access to and efficient use of the data (Figure 2).

The PigQTLdb web interface is designed to be easily accessible by search and / or browse. Each searched or browsed result is again dynamically linked for further search or browse. In this way users can quickly traverse through the information from the PigQTLdb and related databases (Figure 2). The details of a QTL can be summarized (Figure 3) or analyzed in following ways:

◆ Graphical display of multiple QTLs mapped to a chromosome.

PigQTLdb makes it possible for users to "synthesize" a chromosomal map view with multiple QTLs to narrow down a most promising chromosomal region(s) where the best candidate gene(s) for a trait may reside. In Figure 4 is shown an example of multiple QTLs identified by different studies mapped to pig chromosome 7.

◆ Graphical display of QTLs for a trait that are mapped to multiple chromosomes in the pig genome

From the "Trait details" page, a link is provided to show all QTLs mapped in the pig genome for the trait under examination. In Figure 5 is shown an example for "Last Rib Backfat". Users can follow up QTLs on each chromosome to find more supporting QTL evidence for the QTL on that chromosome.

◆ Dynamic link of pig QTL data with related information from other databases

QTL data in the PigQTLdb can be linked to related information in the public databases. As the PigQTLdb is setup at the NCBI and the data are retrievable from the NCBI via Gene, Map Viewer (Figure 6 and 7), and UniSTS databases. All functionality in those tools are available to find related information on pig QTLs through its flanking markers. For example, the match of marker information to records in UniSTS database allows automatic matching of markers to public sequence data by e-PCR.

In addition, the pig QTL information in the NAGRP Animal Genome server and at the NCBI are cross-referenced to each other for each QTL. This function allows all of the unique information on each site to appear to be integrated "on the same server" via worldwide web.

Figure 3. A snap shot of the PigQTLdb web page, showing how the detailed pig QTL information are displayed, and related information are linked.

Discussion

1. There already exist some public QTL databases or genome databases with QTL components, such as those for the mouse obesity, human and mouse bone density, rat, databases for maize, gramene, grain, etc. These QTL database tools are either unavailable to the general public or built as a component to the specific database therefore difficult to be used as a stand-alone database tool. The feasibility to further develop them to include specific functionality that we need is limited.

Figure 4. A snap shot of the chromosomal QTL map view showing multiple QTLs mapped to pig chromosome 7 by "QTL Mapper" tool within the PigQTLdb. Each marker along the chromosome is hyper-linked to the USDA-MARC pig map hosted at NCBI. The QTL names are also hyper-linked to detailed QTL information pages as shown in Figure 3.



Figure 5. A snap shot of QTL map view showing QTLs for "Last Rib Backfat" mapped to multiple chromosomes in the pig genome. The QTL names are hyper-linked to detailed QTL information pages as shown in Figure 3. Each chromosome name is also linked to a graph with all QTLs mapped to that chromosome, as shown in Figure 4.

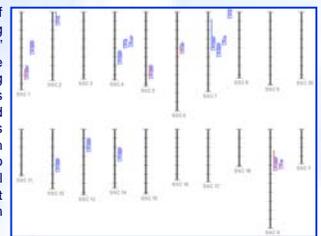
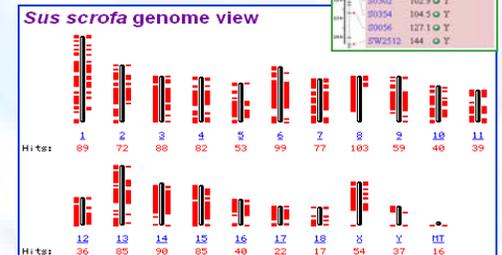


Figure 6. Pig chromosome 1 as shown on the NCBI Map Viewer with QTL information displayed in red bars. Since the Map Viewer is a scalable graphic tool, one can zoom-in/ zoom-out with map tool options on the same page (not shown) to find more specifics.



Figure 7. A snap shot showing the Pig QTL on NCBI Map Viewer. The "Hits" number indicate the number of key word "pig QTL" matches. Each chromosome links to chromosome details in Map Viewer (Figure 6).



Discussion (cont.)

While we realize there has been a lack of a "universal" tool for everybody, we also realize that not everyone makes their database design, database schema and codes easily accessible by the public. By developing the pig QTL database, it is also our goal to make every effort to minimize the redundant work by sharing the tool with the public. As far as the database schema is concerned, the PigQTLdb tools and database structure can be used for other livestock species as is.

2. We introduced the concept of ontology to manage diverse pig production traits for a controlled trait vocabulary management and cross-experiment QTL comparison. We realize this is only a first step towards a full expanded Trait Ontology and we plan to continue our work on making it a more generally acceptable resource / tool for the community.

3. In developing a pig QTL database, we did not develop a static QTL information storage place, but we tried to provide a generic tool for continuing QTL data update, analysis, and comparison. The way how the QTL data in the NAGRP Genome Server and at the NCBI are cross-referenced provide a good example for genome information integration with distributed databases. Our vision is to make it part of the integrated functional genomic information resources for pigs, and for other agricultural species at large.

Acknowledgements

The authors wish to thank Drs. Jack Dekkers from Iowa State University, Andy Law from Roslin Institute (UK) and Lizhen Wang from Sygen International for their useful discussions in the process of the database design. Thanks also due to Drs. David Adelson from Texas A&M University and Per Johnson from Göteborg University (Sweden) for sharing their works on cattle and rat QTL database design, respectively. The data for the template chromosome maps are made possible with data mainly from the USDA-MARC comprehensive pig genome map. This work is supported in part by USDA-NAGRP Pig Genome and Bioinformatics Coordination Programs.

References

Bidel, J. P. and Max F. Rothschild, 2002. Current Status of Quantitative Trait Loci Mapping in Pigs. Pig News and Information, 2002, Vol. 23, No. 2
Rohrer, G. A., L. J. Alexander, Z-L. Hu, T. P. L. Smith, J. W. Keele, and C. W. Beattie. 1996. A comprehensive map of the porcine genome. Genome Research 6:371-391.